

【特許請求の範囲】

【請求項 1】

ユーザから対話のための入力となされると、該入力内容を認識する認識手段と、
該認識手段による認識に応じて、予め記憶された対話対象語の中からユーザに向けた対話語を選択する選択手段と、

該選択手段によって選択された対話語を出力する出力手段と、

を備え、ユーザとの間で対話を行う対話システムであって、さらに、

前記認識手段による認識に基づき、対話上の不整合がある場合に、その不整合を認定する認定手段と、

該認定手段が不整合を認定した際に、予め定める条件に基づいて対話の評価を行い、対話の継続の有無を判定する判定手段と、

を備え、

前記選択手段は、前記判定手段の判定結果に基づき、対話の継続又は終了用に予め設定された対話語を選択することを特徴とする対話システム。

【請求項 2】

ユーザから音声入力となされると、該入力された音声を認識する音声認識手段と、

該音声認識手段によって認識された語に応じて、予め記憶された発話対象語の中から発話語を選択する発話語選択手段と、

該発話語選択手段によって選択された語を音声として出力する発話手段と、

を備え、ユーザとの間で対話を行う対話システムであって、さらに、

前記音声認識手段による認識に基づき、対話上の不整合がある場合に、その不整合を認定する認定手段と、

該認定手段が不整合を認定した際に、予め定める条件に基づいた対話の評価と、その評価に基づく対話継続の有無の判定を行う判定手段と、

を備え、

前記発話語選択手段は、前記判定手段の判定結果に基づき、対話の継続又は終了用に予め設定された発話語を選択することを特徴とする対話システム。

【請求項 3】

前記対話上の不整合は、前記ユーザの発話内容の誤り、前記ユーザとの対話を通じて認定される前記音声認識手段による認識の誤り、又は前記判定手段の評価の誤りから認定されることを特徴とする請求項 2 記載の対話システム。

【請求項 4】

前記予め定める条件は、ユーザプロファイル、確率、音声情報、経過時間、応答速度、及び対話状況のいずれか又はいずれかの組み合わせにより設定されることを特徴とする請求項 2 又は請求項 3 記載の対話システム。

【請求項 5】

前記音声認識手段は、前記音声の認識をロバスト化することを特徴とする請求項 2～4 のいずれかに記載の対話システム。

【請求項 6】

前記ロバスト化において、多様なアクセントや声色を標準化して認識できるようにすることを特徴とする請求項 5 記載の対話システム。

【請求項 7】

前記ロバスト化において、前記ユーザの独り言をカットすることを特徴とする請求項 5 又は請求項 6 記載の対話システム。

【請求項 8】

前記ロバスト化において、応答速度を速くすることを特徴とする請求項 5～7 のいずれかに記載の対話システム。

【請求項 9】

前記判定手段は、予め定める条件により、ユーザの誤りを許す／許さない、譲歩する／譲歩しない、又は問い返すを使い分けることを特徴とする請求項 2～8 のいずれかに記載の

10

20

30

40

50

対話システム。

【請求項 10】

前記発話語選択手段は、予め定める条件により、回答パターンを変化させるように前記発話語を選択することを特徴とする請求項 2～9 のいずれかに記載の対話システム。

【請求項 11】

前記発話語選択手段は、予め定める条件により、作為的に誤りを混入させるように前記発話語を選択することを特徴とする請求項 2～10 のいずれかに記載の対話システム。

【請求項 12】

ユーザから音声入力となされると、該入力された音声を認識する音声認識手段と、
該音声認識手段によって認識された語の最後の音を頭字とする語を、予め記憶された発話対象語の中から選択する発話語選択手段と、
該発話語選択手段によって選択された語を音声として出力する発話手段と、
を備え、ユーザとの間でしりとりを行う対話型しりとりシステムであって、さらに、
前記音声認識手段による認識に基づき、対話上の不整合がある場合に、その不整合を認定する認定手段と、
該認定手段が不整合を認定した際に、予め定める条件に基づいた対話の評価と、その評価に基づく対話継続の有無の判定を行う判定手段と、
を備え、
前記発話語選択手段は、前記判定手段の判定結果に基づき、しりとりの継続又は終了用に
予め設定された発話語を選択することを特徴とする対話型しりとりシステム。 10

【請求項 13】

前記対話上の不整合は、前記ユーザの発話内容の誤り、前記ユーザとの対話を通じて認定される前記音声認識手段による認識の誤り、又は前記判定手段の評価の誤りから認定されることを特徴とする請求項 12 記載の対話型しりとりシステム。

【請求項 14】

前記予め定める条件は、ユーザプロファイル、確率、音声情報、経過時間、応答速度、及び対話状況のいずれか又はいずれかの組み合わせにより設定されることを特徴とする請求項 12 又は請求項 13 記載の対話型しりとりシステム。

【請求項 15】

前記判定手段は、予め定める条件により、ユーザの誤りを許す／許さない、譲歩する／譲歩しない、又は問い返すを使い分けることを特徴とする請求項 12～14 のいずれかに記載の対話型しりとりシステム。 30

【請求項 16】

前記発話語選択手段は、予め定める条件により、前記発話対象語の中から故意に誤りである語を選択し、
前記発話手段が、該誤りである語を発話するとともに、システムの負けを宣言することを特徴とする請求項 12～15 のいずれかに記載の対話型しりとりシステム。

【請求項 17】

請求項 12～16 のいずれかに記載の対話型しりとりシステムにおいて、さらに、しりとりの回答頻度の高い語彙を連ねたパスを予め格納したデータベースを備え、
前記発話語選択手段は、予め定める条件により、前記データベースを参照して前記回答頻度の高い語彙を導くための発話語を選択することを特徴とする対話型しりとりシステム。 40

【請求項 18】

請求項 12～17 のいずれかに記載の対話型しりとりシステムにおいて、さらに、前記ユーザの回答を予測する回答予測手段を備え、
前記発話語選択手段が、該回答予測手段による予測結果に基づいて、前記発話対象語の中から発話語を選択することを特徴とする対話型しりとりシステム。

【請求項 19】

請求項 12～18 のいずれかに記載の対話型しりとりシステムにおいて、さらに計時手段を備え、

前記発話語選択手段は、該計時手段が当該システム側の発話から予め定める経過時間を計時したときに、ユーザをしりとりでの解答に導くためのヒントを選択することを特徴とする対話型しりとりシステム。

【請求項 20】

請求項 12～19 のいずれかに記載の対話型しりとりシステムにおいて、さらに、予め定める条件により、しりとりでの難易度を設定する難易度設定手段を備え、前記発話語選択手段は、該難易度設定手段により設定された難易度に基づいて、前記発話対象語の中から発話語を選択することを特徴とする対話型しりとりシステム。

【請求項 21】

請求項 1～20 のいずれかに記載の対話システム又は対話型しりとりシステムの前記各手段としてコンピュータを機能させるためのプログラム。 10

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、システム—ユーザ間で対話を行うための対話システム及び対話型しりとりシステムに関する。

【0002】

【従来の技術】

従来より、例えばカーナビゲーションシステムにおいてレストラン等の目的地の位置情報を問い合わせたりする情報検索のための対話システム、たわいのない会話をしてユーザを楽しませる娯楽用の対話システム、或いは、しりとり等のゲームに応用した対話システム等、システムとユーザとの対話を可能にした対話システムが知られている。 20

【0003】

【発明が解決しようとする課題】

しかしながら、従来の対話システムでは、以下のような問題点があった。

1) システムにユーザが正しく入力できる場合、認識語彙さえ十分にあればシステムはいつかなるときでも誰に対しても同一内容の対話を展開する。このため、この機械的な対話の単調さにユーザが飽きてしまう。

【0004】

2) 音声で機械とやりとりを行う場合、双方の誤認識に起因して対話が継続されなくなり、対話の楽しさが損なわれることがある。 30

また、従来の対話型しりとりシステムでは、以下のような問題点があった。

1) システムにユーザが正しく入力できる場合、語彙さえ十分にあればシステムは決して負けることなくしりとりを継続する。このため、ユーザがいつも負けることになり気分を害する。

【0005】

2) 特に音声で機械とやりとりを行う場合、双方の誤認識に起因してしりとりが継続されなくなり、楽しさが損なわれることがある。

本発明は、こうした問題に鑑みなされたものであり、たとえユーザ及び対話システムの一方又は双方が誤認識をするような場合でも、ユーザの感情を特に損ねることなく対話やしりとりを継続し、適宜終了できる対話システムを提供することを目的とする。 40

【0006】

【課題を解決するための手段】

上記課題に鑑み、請求項 1 記載の対話システムにおいては、ユーザから対話のための入力となされると、認識手段がこの入力内容を認識する。そして、選択手段が認識手段による認識に応じて、予め記憶された対話対象語の中からユーザに向けた対話語を選択し、出力手段が選択手段によって選択された対話語を出力することで、ユーザとの間で対話を行うものである。

【0007】

そしてさらに、認定手段が、認識手段による認識に基づき、対話上の不整合がある場合に 50

その不整合を認定し、その不整合が認定された際に、判定手段が、予め定める条件に基づいて対話の評価を行い、対話の継続の有無を判定する。そして、上記選択手段は、判定手段の判定結果に基づき、対話の継続又は終了用に予め設定された対話語を選択する。

【0008】

ここでいう「ユーザの入力」は、例えばマイク等を介した音声入力でもよいし、マウスやキーボード等を介した入力であってもよい。また、「出力手段による出力」もスピーカ等を介した音声出力でもよいし、ディスプレイ画面に表示する画像出力等であってもよい。

【0009】

特に音声による入出力によりシステムとユーザ（ユーザ）との対話を実現する場合には、請求項2に記載のようになる。

すなわち、音声認識手段が、ユーザから音声入力となされると、この入力された音声を認識する。そして、発話語選択手段が、この音声認識手段によって認識された語に応じて予め記憶された発話対象語の中から発話語を選択し、発話手段が、この発話語選択手段によって選択された語を音声として出力することでユーザとの間で対話を行う。尚、ここでいう「発話語」とは、上記「対話語」において音声にて出力するものを意味する。

【0010】

そしてさらに、認定手段が、上記音声認識手段による認識に基づき、対話上の不整合がある場合にその不整合を認定し、その不整合が認定された際に、判定手段が、予め定める条件に基づいた対話の評価と、その評価に基づく対話継続の有無の判定を行う。そして、上記発話語選択手段は、この判定手段の判定結果に基づき、対話の継続又は終了用に予め設定された発話語を選択する。

【0011】

かかる構成によれば、音声認識手段によりユーザの音声は正常に認識され、その対話が整合している場合には、そのまま対話が継続される。一方、ユーザの発話内容との関係で対話がかみ合わない、矛盾する、或いは誤っている等により不整合であると認定されたとしても、直ちにユーザ側の発話内容が否定されるわけではない。つまり、実際の正誤に必ずしも拘束されず、予め定める条件に基づいて対話が評価され、戦略上ユーザとの対話を継続するのがよいと判断された場合には、対話が継続される。この場合、対話が不整合と認定された場合に引き続き対話を継続させる際に用いる発話語のバリエーションが予め設定されており、発話語選択手段が、その中から適切な発話語を適宜選択する。

【0012】

従って、ユーザの発言に誤りがあったとしても、それをみだりに否定してその心証を損ねることなく、一定条件の下対話を継続させ、適宜終了させることができる。

この「対話上の不整合」は、例えば請求項3に記載のように、ユーザの発話内容の誤り（つまり人間の誤り）、ユーザとの対話を通じて認定される音声認識手段による認識の誤り、又は判定手段の評価の誤り（つまりシステムの誤り）から認定される。つまり後に詳述するが、この「人間の誤り」には、人間の聞き誤りと人間の回答誤りがあり、「システムの誤り」には、ユーザの発話に対する誤認識と、正常な認識後の発話内容の評価の誤りに基づく発話語の選択誤りがある。尚、その後者のシステム側の誤りについては、システム自身が積極的に認定できるものではないため、予め定める条件の下、ユーザとの対話を通じてシステム側が納得する又は妥協する等により認定する。

【0013】

また、上記「予め定める条件」は、例えば請求項4に記載のように、ユーザプロファイル、確率、音声情報、経過時間、応答速度、及び対話状況のいずれか又は組み合わせにより設定することができる。

例えばユーザプロファイルによるユーザの年齢、性格、地位等の属性に基づき、対話の有効性を厳しく判断したり、逆に緩やかに判断することができる。また、確率的に何回に一回はユーザ側の誤りを許容するといったようにしたり、例えばユーザの口調が強い等の音声情報によって、ユーザの感情が高ぶっている、怒りを感じている等を認定し、緩やかな判断をするようにしてもよい。また、対話を開始してからの経過時間が一定時間を超える

10

20

30

40

50

と対話を終了させる方向に発話語を選択するようにしたり、応答速度により検知されるユーザの焦りにより判断を緩やかにしたりもできる。さらに、渋滞時等の対話状況（状況プロファイル）により対話内容を変更したりすることもできる。尚、これはあくまで一例にすぎず、上記「予め定める条件」には、様々なバリエーションやその組み合わせが考えられる。そして、そのバリエーションによりユーザにシステムとの対話を一層面白く感じるさせることができるのである。

【0014】

つまり、上記対話の評価には、ユーザの誤りやユーザとの対話によるシステム側の誤りの認識、或いは、正誤によらない戦略上の問題等が勘案される。その結果、ユーザ側の認識・思考・発話及びシステム側の認識の各々に誤りが存在しうる場合でも、ユーザの心証を特に損ねずに対話を継続させ、適宜終了させることができるのである。

【0015】

また、ユーザによっては同じ語彙に対する発音等が異なる場合がある。このため、請求項5に記載のように、音声認識手段が音声の認識をロバスト化して、ユーザの発話に柔軟に対応できるようにするのがよい。

その場合、請求項6に記載のように、ロバスト化において、多様なアクセントや声色を標準化して認識できるようにするのが好ましい。これは、例えばいわゆる不特定他者音声認識によって実現することができる。

【0016】

また、請求項7に記載のように、ロバスト化において、ユーザの独り言をカットするようにすると、対話の誤認を防止することができて好ましい。これらの独り言については予め認識辞書に登録しておき、ユーザの発話からその独り言を認識した上で発話内容から削除することにより実現できる。

【0017】

また、請求項8に記載のように、ロバスト化において、応答速度を速くして処理時間を短くすることにより、円滑な対話を実現することができる。その際、対話の無音判定時間を短くすることで応答速度を速くすることができるが、この無音判定時間を短くすれば文の切れ目が判定困難になる。そこで、対話内容に応じた無音判定時間の設定が必要になる。また、話者特性（ゆっくりしゃべる、早口、など）に応じた応答速度（発話速度）の設定も有効である。

【0018】

さらに、請求項9に記載のように、上記判定手段は、予め定める条件により、ユーザの誤りを許す／許さない、譲歩する／譲歩しない、又は問い返すを使い分けるようにすると、対話のバリエーションが豊かになる。

また、請求項10に記載のように、発話語選択手段は、予め定める条件により、回答パターンを変化させるように発話語を選択するとよい。

【0019】

つまり、後述する実施例でも述べるように、同じ発話内容であっても、ユーザによって或いは状況に応じて適宜言い方（表現）を変えることで、対話のおもしろさを向上させることができる。

さらに、請求項11に記載のように、発話語選択手段が予め定める条件により、作為的に誤りを混入させるように発話語を選択するようにすると、人間らしさが感じられ、ユーザに親しみが感じられるようになる。

【0020】

以上のような対話システムは、以下に述べる対話型しりとりシステムとして構成することができる。

すなわち、請求項12に記載の対話型しりとりシステムでは、音声認識手段が、ユーザから音声入力となされると、この入力された音声認識する。そして、発話語選択手段が、この音声認識手段によって認識された語の最後の音を頭字とする語を、予め記憶された発話対象語の中から選択し、発話手段が、この発話語選択手段によって選択された語を音

10

20

30

40

50

声として出力することで、ユーザとの間でしりとりを行う。

【0021】

そしてさらに、認定手段が、上記音声認識手段による認識に基づき、対話上の不整合がある場合に、その不整合を認定し、その不整合が認定された際に、判定手段が、予め定める条件に基づいて対話の評価を行い、しり通りの継続の有無を判定する。そして、上記発話語選択手段は、この判定手段の判定結果に基づき、しり通りの継続又は終了用に予め設定された発話語を選択する。

【0022】

かかる構成によれば、音声認識手段によりユーザの音声は正常に認識され、そのしりとりゲームが正しく進行している場合には、次の発話処理を行い、そのまましりとりが継続される。一方、ユーザの発話内容との関係でしりとりとして不整合であると認定されたとしても、直ちにユーザ側の負けが宣言されるわけではない。つまり、実際の正誤に必ずしも拘束されず、予め定める条件に基づいてしり通りの内容及びその間のユーザとの対話が評価され、戦略上しりとりを継続するのがよいと判断された場合には継続される。従って、ユーザの負けを宣言してその機嫌をみだりに損ねることなくしりとりを継続させ、適宜終了させることができる。

【0023】

この「対話上の不整合」は、例えば請求項13に記載のように、上記と同様にユーザの発話内容の誤り（つまり人間の誤り）、ユーザとの対話を通じて認定される音声認識手段による認識の誤り、又は判定手段の評価の誤り（つまりシステムの誤り）から認定される。

【0024】

また、上記「予め定める条件」は、請求項14に記載のように、上記と同様にユーザプロフィール、確率、音声情報、経過時間、応答速度、及び対話状況のいずれか又は組み合わせにより設定することができる。

さらに、請求項15に記載のように、上記判定手段は、予め定める条件により、ユーザの誤りを許す／許さない、譲歩する／譲歩しない、又は問い返すを使い分けるようにすると、上述のように対話のバリエーションが豊かになる。

【0025】

また、請求項16に記載のように、上記発話語選択手段が、予め定める条件により、発話対象語の中から故意に誤りである語を選択し、発話手段が、この誤りである語を発話するとともにシステムの負けを宣言するようにすれば、ユーザに対して親近感を形成させることができる。

【0026】

また、請求項17に記載のように、データベースにしり通りの回答頻度の高い語彙を連ねたパスを予め格納し、発話語選択手段が、予め定める条件により、データベースを参照して回答頻度の高い語彙を導くための発話語を選択するようにするのもよい。尚、ここでいう「回答頻度」は、個々のユーザの回答頻度であってもよいし、当該システムで過去にしりとりゲームを行った者の全体の回答頻度であってもよい。このようにすれば、ユーザにとって語彙が思いつきやすくスムーズにしりとりを継続することができるため、状況に応じてユーザの気分を良くすることができる。

【0027】

さらに請求項18に記載のように、さらに、回答予測手段がユーザの回答を予測し、発話語選択手段が、この回答予測手段による予測結果に基づいて、発話対象語の中から発話語を選択するようにするのもよい。例えば、それまでのしりとり対話の履歴を参照して、xで始まる単語の発話が少ない場合（つまりユーザが思いつき難い場合）、難易度を高く設定する場合にはこのxを語尾とする単語を選択し、逆に難易度を低く設定する場合には、当該xを語尾とする単語を選択しないようにすることができる。

【0028】

かかる構成によれば、ユーザが回答を思いつきやすい単語と、思いつきにくい単語とを戦略的に用いることで、一層しりとり対話の面白さを高めることができる。

10

20

30

40

50

また、請求項 19 に記載のように、計時手段が、システム側の発話からの時間を計時し、発話語選択手段が、この計時手段が予め定める経過時間を計時したときに、ユーザをしりとりでの解答に導くためのヒントを選択することで、ユーザの負けを減らし、ユーザの気分を良好に維持することができる。

【0029】

さらに、請求項 20 に記載のように、難易度設定手段が、予め定める条件により、しり通りの難易度を設定し、発話語選択手段は、この難易度設定手段により設定された難易度に基づいて、発話対象語の中から発話語を選択するようにしてもよい。この難易度は、回答までの時間の設定、ジャンルの限定、ヒントの有無、上記履歴に基づく語彙の選択等により、設定することができる。

【0030】

かかる構成により、ユーザのレベルを考慮した難易度のしりとりゲームを実現することができる。尚、このような対話システム又は対話型しりとりシステムの各手段をコンピュータにて実現する機能は、例えば、コンピュータ側で起動するプログラムとして備えることができる（請求項 21）。このようなプログラムの場合、例えば、FD、MO、DVD、CD-ROM、ハードディスク等のコンピュータ読取可能な記録媒体に記録し、必要に応じてコンピュータにロードして起動することにより用いることができる。その他、ROM やバックアップ RAM をコンピュータ読取可能な記録媒体としてプログラムを記録しておき、この ROM 或いはバックアップ RAM をコンピュータに組み込んでもよい。

【0031】

【発明の実施の形態】

以下、本発明の実施の一形態を図面と共に説明する。尚、本実施形態は、本発明の対話システムを対話型しりとりシステムとして構成したものであり、図 1 はその対話型しりとりシステムの全体構成を表すブロック図である。

【0032】

同図に示すように、対話型しりとりシステム 1 は、音声認識部 10、対話管理部 20、しりとり処理部 30、及び音声合成部 40 等から構成されている。

そして、ユーザの発話音声は、図示しないマイクを介してまず音声認識部 10 に入力される。音声認識部 10 は、しり通りの語彙や対話に必要な語彙が格納された辞書 11 を参照して、この発話音声の内容を認識し、この認識結果を対話管理部 20 に出力する。対話管理部 20 は、この認識結果から対話の進行を管理し、しりとり処理部 30 やユーザプロフィール 21、状況プロフィール 22、及び発話データベース 23 を参照して応答内容を生

成し、音声合成部 40 に出力する。

【0033】

ユーザプロフィール 21 は、ユーザの性別、年齢、嗜好、性格、職業、地位等といった登録ユーザの属性を示すデータを格納したデータベースである。また、状況プロフィール 22 は、本システムが現在どこでどのように使用されているかを表す使用環境等を格納したデータベースである。さらに、発話データベース 23 は、ユーザプロフィール 21 に格納された個々のユーザの属性に対応して、発話の種類を格納したデータベースである。例えば、同じ対話語であっても、ユーザが地位の高い大人の属性を有する場合に敬語を駆使した対話語を対応させたり、ユーザが若者や子供等の場合にフレンドリーな対話語を対応させたりする等、種々のバリエーションでの発話の種類が格納されている。

【0034】

しりとり処理部 30 は、しりとり専用の語彙を格納した語彙データベース 31 を備え、後述するしり通りの正誤判断、しり通りの回答予測、しり通りの応答生成を、この語彙データベース 31 を参照して行い、最終的に総合的なしり通りの勝負の判断を行う。図 2 にしりとり処理部 30 の機能の一例を示す。

【0035】

同図に示すように、しりとり処理部 30 では、現在の時刻、しりとり開始からの経過時間

10

20

30

40

50

、しりとり各応答時間等を逐次検出する一方で、ユーザプロフィール、状況プロフィール、時事情報、設定入力等に基づいて、現在のユーザの発話に対して戦略上適切な評価を行い、選択した発話語（判定結果）を出力する。

【0036】

ここで、設定入力とは、例えば意図的に誤りを入れる／入れない、ユーザの好みの言葉を選択する／しない、嫌いな言葉を選択する／しない、回答しやすい言葉を選択する／しないといった具合に、ゲームの面白さを高める上での種々の設定であり、しりとりゲームの開始前にユーザ又は第三者により予め設定される。

【0037】

また、後述する音声のロバスト化のために、ユーザの発話音声の特徴を記録しておいたりする。また、判定結果について複数の適合値が見つかったような場合等には、乱数入力によりそのいずれかを選択するようにすることもできる。

そして、図1に戻り、音声合成部40は、対話管理部20にて生成された応答内容を音声に変換し、図示しないスピーカを介して発話し、これをユーザに伝える。

【1. 全体の流れ】

次に、本対話型しりとりシステム1が実行するしりとりゲームに関する処理について、図3に示すフローチャートに基づいて説明する。

【0038】

まず、しり通りの開始に際してユーザとの間で導入対話を行う（S110）。この導入対話では、システム側から例えば「こんにちは」、「これからしりとりを始めるよ」、「君のお名前は？」、「どこから来たの？」、「何歳？」、「どんな仕事をやってるの？」といった発話が行われ、これに対するユーザの発話を認識してユーザプロフィール21や状況プロフィール22を参照し（S120）、ユーザの認証や属性判別等をして発話データベース23における発話の種類を選択するとともに、しり通りのジャンル設定や難易度の設定等を行う。

【0039】

そして、本システム側からしりとりを開始する場合にはまず、しりとり開始からの経過時間をチェックした後（S130）、語彙データベース31から最初の語彙として適切な一語を選択して発話する（S140、150）。この最初の語彙としては、例えば、ユーザプロフィール21を基にユーザが容易に思いついて比較的解答しやすいものを選択し、スムーズにしりとりを開始することが考えられる。

【0040】

次に、ユーザからこれに続く適切と思われるしり通りの単語が発話されると、このユーザ発話を音声認識する（S160）。そして、この認識した単語がしりとりとして正しいかどうかを評価する（S170）。

そして、しり通りのルールに照らして正しいと判断された場合には（S180：YES）、S130に戻り、経過時間をチェックした後、次の応答語彙を生成してしりとりを継続する（S140～S170）。一方、正しくないと判断された場合には（S180：NO）、総合的な勝負の判断を行う（S190）。この総合的な判断は、後述する「2. 音声対話における誤りの分類」等に基づき、経過時間情報、ユーザプロフィール、時間帯、本システムの使用環境（状況プロフィール）等により判断する。

【0041】

例えば、ユーザが偉い人であれば適当なところでわざとシステムの負けを宣言してゲームを終了したり、ユーザがゲームにのめり込むような子供である場合に厳しく勝敗を決定したりすることが考えられる。また、ユーザの主張や反論によりそのままゲームを継続させたり、しりとりゲーム開始からの一定時間経過によりシステムの負けを宣言してゲームを終了するようにしてもよい。

【0042】

そして、上記総合的な判断により、ユーザの勝ちと判断されれば（S200）、システムがユーザの勝ちを宣言する発話を行う（S210）。

10

20

30

40

50

一方、システムの勝ちと判断されれば（S 2 0 0）、システムが勝ちを宣言する発話を行う（S 2 2 0）。このとき、ユーザからの主張（反論）があるか否かを確認し（S 2 4 0）、何もなければ（S 2 4 0 : N O）、そのままゲームを終了する。一方、S 2 4 0においてユーザからの主張がある場合には（S 2 4 0 : Y E S）、ユーザとの対話を行う。

【 0 0 4 3 】

そして、S 2 0 0において勝ち負けが決まらない場合（S 2 0 0）、及びS 2 4 0においてユーザとの対話を行った場合には、ユーザの機嫌を損ねないように戦略上わざと負けるかどうかを判断する（S 2 3 0）。そして、わざと負けると判断した場合には（S 2 3 0 : Y E S）、システムがユーザの勝ちを宣言する発話を行って一連の処理を終了する（S 2 1 0）。一方、わざと負けることをしないと判断した場合には（S 2 3 0 : N O）、S 1 3 0に戻ってそのまましりとりを継続し、次の応答を生成する。

〔 2 . 音声対話における誤りの分類 〕

2 . 1 人間の誤り

しり通りのプロセスにおいて人間側で生じうる誤りプロセスには下記の2つがある。

（1）「人間の聞き誤り」として、システムの発話を誤認識する場合がある。

（2）「人間の回答あやまり」として、しり通りの応答として誤った答えを作成する場合がある。

2 . 2 システムの誤り

また、しり通りのプロセスにおいてシステム側で生じうる誤りプロセスには下記の2つがある。

（1）「入力語に対するシステムの認識誤り」として、例えば音声認識部 1 0 における誤認識がこれに相当する。これは、ユーザの発話に対するシステムの音声認識に誤りがあるが、その後発話を進める上でのシステムの判断（評価）に誤りがない場合である。

（2）「システムの判断誤り」がある。これは、ユーザの発話に対するシステムの音声認識に誤りがない場合でも、その後発話を進める上でのシステムの判断（評価）に誤りがある場合である。

【 0 0 4 4 】

ただし、プログラムや装置上のバグがなければ、しり通りの論理や個々のタスクについて決められた条件判断についてシステムが誤りを犯すとは考えにくい。このため、基本的には上記の入力語が音声認識で誤認識され、結果としてシステムの判断誤りに“見える”ケースが多いと考えられる。一方で、いくつかある選択肢から状況やユーザプロファイル、確率計算に応じてひとつの戦略を選択する場合が重要である（下記参照）。

（例）例えば、ユーザがしりとりにならない発話をしたと判断した場合、

- 1) ユーザの誤りを許す
- 2) ユーザの誤りをゆるさない
- 3) わざと負ける
- 4) 問い返す

などの対応戦略が考えられる。もしこれらの中からひとつの戦略を選ぶ場合、経過時間やユーザプロファイル、時間帯、使用環境（状況プロファイル：交通渋滞、パーキングエリア他）によって戦略を選択していくことになる。システム側は、これらの戦略を選ぶ際にその選び方がユーザに適しているか否かの判断を誤る可能性がある。例えば、本システムが車両に搭載され、車内でゲームが進行している場合に、使用環境において交通渋滞であると判断した直後に、車が流れ出したような場合が該当する。

〔 3 . 適応化する機能 〕

尚、上述した緒機能は下記のように集約される。

3 . 1 わざと負ける

ゲームの時間（経過時間）が長くなったり、音声認識がうまくいかない、或いは対話のやりとりがうまくいかない場合には、システム側はわざと負けることを選択することができる。

3 . 2 人間のいい直しを許す

10

20

30

40

50

最初に言った回答をシステムが認識した後でも、次の応答を返す以前にユーザが「ちょっとまって」などと割り込んでこれをすばやく取り消し、言い直す場合でも受け付ける。

3. 3 人間のいい誤りを許す

(1) 人間が本当に間違えて言った場合

- a) 指定ジャンル以外の言葉を使った
- b) 本当にしりとりにならない言葉を使った
- c) 考えている最中の独り言を誤認識した

(2) システムが認識した結果、人間の言い誤りと判定した場合

3. 4 回答パターンを微妙に変化させる

システムが勝った場合或いはユーザが返答に困っている場合などに、単に「僕の勝ちだね」10
というだけではなく、「もうすこしががんばりましょう」などと励ましの言葉を入れたりする。「僕の負けだね」は「あー負けちゃった」などと適宜言い方を変えることが自然さを感じさせる上で有効である。このためには例えば、システムが負けた場合の発話クラス（同じ内容の発話）として、

- 1) 僕の負けだね
- 2) 私の負けです、お強いですね
- 3) 君の勝ちだね
- 4) うーん、もう思いつかないな、えーと、あー時間がない
- 5) あー負けちゃった
- 6) x x x、あ、これは前に言ったね、うーんギブアップ。

20

などの中からユーザのプロファイル（性別、年齢層、職業など）やゲームの経過時間などから適宜選択して違う回答を出せるようにする。

3. 5 ヒントを出す

ユーザが回答に困っている場合、つまり、例えば回答待ちの時間が一定時間を超過した場合や、「えーと」、「うーん」等の発話を認識した場合、或いは複数回間違った回答をしようとした場合には、例えば下記のようなヒントを出す。

<動物しりとり例>

- 「魚や鳥の名前でもいいよ」
- 「”み”のつく鳥がいるよね」
- 「”す”のつく魚は何だったかな」

30

尚、この機能は、しりとりを開始する際の後述する難易度の選定や回答予測機能、回答データベースにも関連する。

[4. 認識のロバスト化]

4. 1 多様なアクセントや声色、話速への対応

同じ単語でも人によってさまざまな言い方をするため、例えば「みみずく」をゆっくり「みーみーずく！」と発話された場合でも認識できるように考えられる言い方を予め認識辞書（辞書11）中に記述しておく。アクセントや声色にも影響されるが、現時点ではその手段はいわゆる不特定他者音声認識によって吸収する。

4. 2 独り言をカットする

ユーザが何をいっていいかわからない場合、その間の独り言が認識語にはいってしまう。40
この場合、

- 「えーと何があるかな」
- 「うーんわからないな」

などの考えられる独り言を予め認識辞書に登録しておき、認識した上で発話内容から削除する。

4. 3 応答速度を早くする

発話開始スイッチ（発話の際に雑音や余計な音声が入らないようにするために、発話開始時に押すスイッチ）の不要な音声認識の場合、通常、発話音声から無音区間を検出することで文の切れ目を判定することが多い。ここでその無音判定時間を短くすれば、システムの処理時間は短くなるが、逆に文の切れ目が判定困難になる。そこで、対話内容に応じた

50

無音判定時間の設定が必要になる。また、話者特性（ゆっくりしゃべる、早口、など）に応じた設定も有効である。

[5. その他の機能]

5. 1. 履歴しりとり

履歴を言うしりとりもできるようにする。

(例) 「りんご」

『りんご、ゴリラ』

「りんご、ゴリラ、らっきょう」

『りんご、ゴリラ、らっきょう、うさぎ』

のように、システム及びユーザが、それまでのしりとりの履歴を記憶するしりとりゲーム 10

5. 2 難易度を設定する

下記要素で難易度をつける。

[0 0 4 5]

- ・ 応答時間の限定：10秒以内、30秒以内、1分以内、無制限
- ・ 語彙限定：動物しりとり、たべものしりとり、国名しりとり、他
- ・ ヒントの有無：
- ・ わざと負けるかどうか：

これらは、ユーザの属性に合わせて設定する。

5. 3 しりとりにおける回答のデータベース化 20

しりとりにおいては回答しやすいパスを予めデータベース化しておくことができる。

(例) ・ きつね、ねこ、コアラ、らくだ、だちょう、うさぎ、ぎんぎつね、。。。

[0 0 4 6]

・ りんご、ごま、マカロニ、にんにく、くり、。。。

これにより、ユーザが思いつきやすく解答しやすくなる。

5. 4 しりとりにおける回答の予測

ボキャブラリのネットワーク構造の解析により、発話「りんご」に対して「ご」で始まる単語が予測できるため、その予測単語集合：(ご) = {ごりら、ごま、ごぼう、ごーや、e t c.} をもとに、

1) ヒントを出す

2) 予測単語集合 $P(x)$ に含まれるが、まだ一度も発話されていない単語の数 $N(P(x))$ が、あるしきい値 N_{th} に対して

$$N(P(x)) \leq N_{th}$$

の場合、これは回答を思いつきにくい問題であると判断して、 x で終わるような単語を発話しない。逆に難易度の高い設定であるならば $N(P(x))$ の小さい x で終わるような単語を選んで発話する。

[0 0 4 7]

また、これらの回答の予測結果とユーザプロフィールを組み合わせ、

- ・ ユーザの嫌いな言葉を言わない、ユーザにも言わせない
- ・ ユーザの好きな言葉を選択して応答する、或いはユーザに言わせられるようにする 40
- などが考えられる。逆にユーザにストレスを与えたい場合には、
- ・ ユーザの嫌いな言葉を言わせる、あるいはユーザに言わせる
- ・ ユーザの嫌いな言葉を選択して応答する、あるいはユーザに言わせられるようにする
- などが考えられる。

5. 5 対話におけるキーワードの予測

上記のしりとりにおける語尾ー語頭の関係に対話におけるキーワードの連想関係、意味関係に置き換えれば、ある発話中の単語に対して応答に用いる関連（又は連想）キーワード（キープレーズでもよい）群を特定することができる。例えば、

$KW(りんご) = \{くだもの、みかん、たべもの、赤い、私の好きなくだもの、$

$アップル、玉露、硬い、甘ずっぱい、青い、e t c. \}$

と表現できる。さらに、このキーワード群の個々の単語に対しても同様に、

KW (私の好きなくだもの) = {みかん、キーウィ、いちご、パイナップル、りんご、デザート、食事、おやつ、フルーツ、ビタミン、野菜、肉、魚、おなかがすいた、他}

とこの連関ネットワークを広げることができる。このようにして活性化されたキーワードやキーフレーズを連想ゲームに用いることもできるし、また、対話の話題展開につなげることもできるが、そこにおいてしりとりするときと同様にユーザプロフィール情報を用いてユーザの好きなキーワードあるいはユーザの嫌いなキーワードまたはどちらでもない(特に嗜好性のない)キーワードなどを意図的に選択することができる。

5. 6 意図的に誤りを混入する

わざと聞き間違えたようなふりをして

「ゴリラ」と正しく認識したものを→「えっコアラ?」、「甲羅?」

「りんご」と正しく認識したものを→「えっピンゴ?」

などと取り違えたふりをして聞き返す等して面白さを演出する。

【0048】

尚、本実施形態のように音声入力でなくテキスト入力の場合には、かな漢字変換誤りなどをわざと入れることも考えられる。例えば図2に示すように、ユーザが魚の「すずき」と回答したのを、「鈴木君がどうかしたの?」「鈴木xx大臣のこと?」などと取り違えたふりをして聞き返す等して面白さを演出する。

【実施例】

図4に、対話型しりとりシステム1にて実行されるしりとりにおける条件分岐とシステムの発話戦略の実施例を示す。以下に、各条件における動作内容を説明する。

【1. システムの発話『りんご』に対して人間(ユーザ)が正しく「ゴリラ」と発話する場合】

Case-1a) システムが正しく「ゴリラ」と音声認識し、語の接続が正しいと判断する。これを踏まえて次の語『ラッパ』を発話する。ユーザも接続が正しいという印象をもつ。この場合、しり通りの結果は「未定」となり、しりとりは継続される。

【0049】

Case-1b) システムが「ごま」と誤認識するが、語の接続は正しいと判断する。システムはこれを踏まえて次の語『まり』を発話する。ユーザはおかしいと思うがそのまま続ける。

Case-1c) システムが「ごま」と誤認識するが、語の接続は正しいと判断する。これを踏まえて次の語『まり』を発話する。ユーザはおかしいと思い、自分の勝ち(システムの負け)を宣言する。システムはこれを受け入れ、負けを宣言する。

【0050】

Case-1d) システムが「ごま」と誤認識するが、語の接続は正しいと判断する。しかし、「ごま」はしり通りの経緯から見て2回目の発話であり、既に出ているのでこれはユーザの負けと判断し、システムの勝ちを宣言する。しかし、ユーザがこれに納得せずに自分が正しいことを主張する場合には、しり通りの結果は「未決」となり、上述したユーザとの対話の中でシステムがわざと負けるか否かを判断する。

【0051】

Case-1e) システムが「コレラ」と誤認識し、語の接続がおかしいと判断する。しかし、戦略的に穏便にしりとりを継続すべく、この間違い(本当はユーザは正しい)を許容し、次の語『ラッパ』を発話する。この場合、しり通りの結果は「未定」となり、しりとりは継続される。

【0052】

Case-1f) システムが「コレラ」と誤認識し、語の接続がおかしいと判断する。システムは自分の勝ちを宣言するが、ユーザはその判定をおかしいと思う場合には、抗議の発話をする。一方、ここでユーザが何も抗議しなければそのままユーザの負けとなる。

【0053】

Case-1g) システムが「コレラ」と誤認識し、語の接続がおかしいと判断。シス

10

20

30

40

50

テムは戦略により「そろそろ負けてやろうか」と考えて『ライオン』を発話し、わざと負け、システムの負けを宣言する。

〔2. システム発話『りんご』に対してユーザが「ゴリラ」（2回目）を発話する場合〕

Case-2 a) システムが「ゴリラ」と認識し、その発話が2回目であることからユーザの負けと判断する。システムは自分の勝ちを宣言する。

【0054】

Case-2 b) システムが「ごま」と誤認識し、『まり』と応答する。ユーザは自分の言った言葉「ゴリラ」に対してしりとりになっていないことに気が付かないか或いは気が付いても、システムの間違いを許容し、しりとりを継続する。

Case-2 c) システムが「ごま」と誤認識し、『まり』と応答する。ユーザは自分の言った言葉「ゴリラ」に対してしりとりになっていないことに気が付いて自分の勝ちをシステムに宣言する。システムはこれを素直に受け入れて負けを宣言し、しりとりは終了する。

【0055】

Case-2 d) システムが「ごま」と誤認識し、戦略によりそろそろ負けてやろうと考えて『マリオン』と応答すると同時に自分の負けを宣言する。一方、ユーザは自分の言った言葉「ゴリラ」に対してしりとりになっていないことに気が付いて自分の勝ちを当然と納得する。

【0056】

Case-2 e) システムが「ごま」（2回目）と誤認識し、『2回目だよ。私の勝ち』と言ってユーザの負けを宣言する。一方、ユーザは自分の発話した「ゴリラ」が2回目であることに気づいて自分の負けを納得する。

Case-2 f) システムが「ごま」（2回目）と誤認識し、『"ごま"は2回目だよ。私の勝ち』と言ってユーザの負けを宣言する。一方、ユーザは自分は「ゴリラ」と正しく発話した（実は2回目）ので自分の勝ちを主張する。勝敗は未決だが戦略により必要に応じてシステムが負けを認める。

【0057】

Case-2 g) システムが「コレラ」と誤認識し、『りんご』に対するしりとりになっていないことに気が付くが許容し、『ラッパ』と応答してしりとりを継続する。

Case-2 h) システムが「コレラ」と誤認識し、『りんご』に対するしりとりになっていないと判断して自分の勝ち（ユーザの負け）を宣言する。これに対してユーザは自分のいった「ゴリラ」が2回目であることに気がつかないまま自分の発話がしりとりとして正しいことを主張する。システムがその後「コレラ」ではなく「ゴリラ」と言ったのだと理解できれば「ゴリラ」が2回目であることを説明してユーザを納得させる。理解できなければ、少なくとも自分の勝ちではないと判断してそのまましり通りの継続を選択することもできる。

【0058】

Case-2 i) システムが「コレラ」と誤認識し、『りんご』に対するしりとりになっていないことに気づくが、戦略により「そろそろ負けてやろうか」と考えて『ライオン』を発話してわざと負け、その負けを宣言する。

〔3. システム発話『りんご』に対してユーザが「ドラマ」と発話する場合〕

Case-3 a) システムが「ドラマ」と正しく認識し、しりとりになっていないことに気づいても、戦略により許容してしりとりを継続し、『まり』と発話する。

【0059】

Case-3 b) システムが「ドラマ」と正しく認識し、しりとりになっていないことに気づいて自分の勝ちを宣言する。

Case-3 c) システムが「ドラマ」と正しく認識し、しりとりになっていないことに気づくが、戦略により「そろそろ負けてやろうか」と考えて『マリオン』を発話してわざと負け、その負けを宣言する。

10

20

30

40

50

【0060】

Case-3d) システムが「ゴリラ」と誤認識し、正しいしりとりであると判断して『ラッパ』と応答する。一方、ユーザは「ドラマ」と発話したのに『ラッパ』と返ってきたのでおかしいと感じるが許容し、そのまま継続する。

Case-3e) システムが「ゴリラ」と誤認識し、正しいしりとりであると判断して『ラッパ』と応答する。一方、ユーザは「ドラマ」と発話したのに『ラッパ』と返ってきたのでおかしいと感じ、自分の勝ちを主張する。システムは戦略により負けを認め、その負けを宣言する。

【0061】

Case-3f) システムが「ケラマ」と誤認識し、しりとりになっていないことに気づいても許容してしりとりを継続し、『まり』と発話する。しりとりの結果は「未定」となり、しりとりは継続される。

Case-3g) システムが「ケラマ」と誤認識し、しりとりになっていないことに気づいて自分の勝ちを宣言する。

【0062】

Case-3h) システムが「ケラマ」と誤認識し、しりとりになっていないことに気づくが、戦略により「そろそろ負けてやろうか」と考えて『マリオン』を発話してわざと負け、その負けを宣言する。

【4. システム発話『りんご』に対してユーザが「ゴメス」（未知語）と発話する場合】

Case-4a) システムが「ゴリラ」と誤認識し、正しいしりとりであると判断して『ラッパ』と応答する。一方、ユーザは「ゴメス」と発話したのに『ラッパ』と返ってきたのでおかしいと感じるが許容し、そのまま継続する。

【0063】

Case-4b) システムが「ゴージャス」と誤認識し、正しいしりとりであると判断して『すいか』と応答する。一方、ユーザはシステム内部の誤認識に気づかず、正しいしりとりであると判断してそのまま継続する。

Case-4c) システムが「xxx」（2回目）と誤認識し、ユーザの負けを宣言する。これに対して、ユーザは正しく言ったと主張する。その結果、戦略により継続するか必要に応じてシステムが負けを認める。

【0064】

Case-4d) システムが「こりす」と誤認識し、しりとりになっていないと判断するが許容してしりとりを継続し、『すいか』と発話する。しりとりの結果は「未定」となり、しりとりは継続される。

Case-4e) システムが「こりす」と誤認識し、しりとりになっていないことに気づいて自分の勝ちを宣言する。しかし、ユーザがこれに納得せずに自分が正しいことを主張する場合には、しりとりの結果は「未決」となり、上述したユーザとの対話の中でシステムがわざと負けるか否かを判断する。

【0065】

Case-4f) システムが「こりす」と誤認識し、しりとりになっていないことに気づくが、戦略により「そろそろ負けてやろうか」と考えて『スワン』を発話してわざと負け、その負けを宣言する。

以上に説明したように、本実施形態の対話型しりとりシステム1は、ユーザの発話内容の誤りや、ユーザとの対話を通じて認定される音声認識部10による認識の誤りにより、しりとり対話に不整合が生じたとしても、直ちにユーザ側の発話内容が否定されるわけではない。つまり、実際の正誤に必ずしも拘束されず、総合的な勝負の判断により戦略的に対話が評価され、ユーザとの対話を継続するのがよいと判断された場合には、対話が継続される。従って、ユーザをみだりに否定してその心証を損ねることなく、一定条件の下対話を継続させて適宜終了させることができる。

【0066】

10

20

30

40

50

特に、総合的な勝負の判断は、ユーザプロファイル、確率、音声情報、経過時間、応答速度、及び対話状況等により、様々なバリエーションで予め設定することができ、そのバリエーションによりユーザがシステムとの対話を一層面白く感じることができる。

【0067】

尚、本実施形態において、音声認識部10が音声認識手段に該当し、対話管理部20及びしりとり処理部30が、発話語選択手段、認定手段、判定手段、回答予測手段、計時手段、及び難易度設定手段に該当し、音声合成部40が発話手段に該当する。

【0068】

以上、本発明の一実施形態について説明したが、本発明の実施の形態は、上記に示したものに何ら限定されることなく、本発明の技術的範囲に属する限り種々の形態をとり得ること

10

とはいうまでもない。
例えば、上記実施形態では、本発明の対話システムを対話型しりとりシステムとして実現した例を示したが、しりとりシステム以外の対話システムであってもよいことはもちろんである。

【0069】

また、上記実施形態では対話を音声により行う例を示したが、テキスト入力により対話を

【図面の簡単な説明】

【図1】本発明の一実施形態にかかる対話型しりとりシステムの概略構成を表すブロック図である。

【図2】対話型しりとりシステムの処理機能の例を表す説明図である。

【図3】対話型しりとりシステムにおける処理を表すフローチャートである。

【図4】しりとり対話における実施例を表す説明図である。

【符号の説明】

1・・・対話型しりとりシステム、 10・・・音声認識部、 11・・・辞書、 20・・・対話管理部、 21・・・ユーザプロファイル、 21・・・ユーザプロファイル、 22・・・発話データベース、 30・・・しりとり処理部、 31・・・語彙データベース、 40・・・音声合成部

20

フロントページの続き(51) Int. Cl. ⁷

G 1 0 L 15/10

G 1 0 L 15/22

F I

G 1 0 L 3/00 5 7 1 U

G 1 0 L 3/00 5 5 1 H

G 1 0 L 3/00 5 3 1 J

テーマコード (参考)